Probability Theory Overview and Analysis of Randomized Algorithms

Analysis of Algorithms

Prepared by
John Reif, Ph.D.
Probability Theory Topics

a) Random Variables: Binomial and Geometric

b) Useful Probabilistic Bounds and Inequalities
Readings

• Main Reading Selections:
  – CLR, Chapter 5 and Appendix C
Probability Measures

- A probability measure (Prob) is a mapping from a set of events to the reals such that:
  1) For any event A
     \[ 0 \leq \text{Prob}(A) \leq 1 \]
  2) \( \text{Prob} \) (all possible events) = 1
  3) If A, B are mutually exclusive events, then
     \[ \text{Prob}(A \cup B) = \text{Prob} (A) + \text{Prob} (B) \]
Conditional Probability

- Define

\[ \text{Prob}(A \mid B) = \frac{\text{Prob}(A \land B)}{\text{Prob}(B)} \]

for \( \text{Prob}(B) > 0 \)
Bayes’ Theorem

- If $A_1, ..., A_n$ are mutually exclusive and contain all events then

$$\text{Prob}(A_i \mid B) = \frac{P_i}{\sum_{j=1}^{n} P_j}$$

where $P_j = \text{Prob}(B \mid A_j) \cdot \text{Prob}(A_j)$
Random Variable A
(Over Real Numbers)

- Density Function

\[ f_A(x) = \text{Prob}(A=x) \]
Random Variable A (cont’ d)

- Prob Distribution Function

\[ F_A(x) = \text{Prob}(A \leq x) = \int_{-\infty}^{x} f_A(x) \, dx \]
Random Variable A (cont’d)

- If for Random Variables $A, B$
  \[ \forall x \quad F_A(x) \leq F_B(x) \]

- Then “$A$ upper bounds $B$” and “$B$ lower bounds $A$”

\[ F_A(x) = \text{Prob} (A \leq x) \]
\[ F_B(x) = \text{Prob} (B \leq x) \]
Expectation of Random Variable $A$

\[ E(A) = \overline{A} = \int_{-\infty}^{\infty} x f_A(x) \, dx \]

- $\overline{A}$ is also called “average of $A$” and “mean of $A$” = $\mu_A$
Variance of Random Variable A

\[ \sigma^2_A = (A - \bar{A})^2 = A^2 - (\bar{A})^2 \]

where 2nd moment \( A^2 = \int_{-\infty}^{\infty} x^2 \, f_A(x) \, dx \)
Variance of Random Variable A (cont’d)
Discrete Random Variable A

\[ f_A(x) = \text{prob} \ (A=x) \]
Discrete Random Variable A (cont’d)

$F_A(x) = \sum_{i=0}^{x} f_A(i) = \text{prob} (A \leq x)$
Discrete Random Variable $A$ Over Nonnegative Numbers

- Expectation

$$E(A) = \bar{A} = \sum_{x=0}^{\infty} x f_A(x)$$
Pair-Wise Independent Random Variables

- A,B independent if

\[ \text{Prob}(A \land B) = \text{Prob}(A) \times \text{Prob}(B) \]

- Equivalent definition of independence

\[
\begin{align*}
    f_{A \land B}(x) &= f_A(x) \cdot f_B(x) \\
    M_{A \land B}(s) &= M_A(s) \cdot M_B(s) \\
    G_{A \land B}(z) &= G_A(z) \cdot G_B(z)
\end{align*}
\]
Bounding Numbers of Permutations

- \( n! = n \times (n-1) \times 2 \times 1 \)
  \( = \) number of permutations of \( n \) objects

- Stirling’s formula
  \( n! = f(n) \times (1+o(1)) \), where
  \[ f(n) = n^n e^{-n} \sqrt{2\pi n} \]
Bounding Numbers of Permutations (cont’d)

• Note
  – Tighter bound

\[
f(n) \frac{1}{e^{(12n+1)}} < n! < f(n) \frac{1}{e^{12n}}
\]

\[
\frac{n!}{(n-k)!} = \text{number of permutations of n objects taken k at a time}
\]
Bounding Numbers of Combinations

\[
\binom{n}{k} = \frac{n!}{k! (n-k)!}
\]

= number of (unordered) combinations of \(n\) objects taken \(k\) at a time

- Bounds (due to Erdos & Spencer, p. 18)

\[
\binom{n}{k} \sim n^k e^{-\frac{k^2}{2n}} - \frac{k^3}{6n^2} (1 - o(1))
\]

for \(k = o\left(\frac{3}{n^4}\right)\)
Bernoulli Variable

- $A_i$ is 1 with prob $P$ and 0 with prob $1-P$

Binomial Variable
- $B$ is sum of $n$ independent Bernoulli variables $A_i$ each with some probability $p$

```
procedure BINOMIAL with parameters n,p
begin
  B ← 0
  for i=1 to n do
    with probability $P$ do B ← B+1
  output
end
```
B is Binomial Variable with Parameters n, p

*mean* \[ \mu = np \]

*variance* \[ \sigma^2 = np(1-p) \]
B is Binomial Variable with Parameters n, p (cont’d)

density fn = \text{Prob}(B=x) = \binom{n}{x} p^x (1-p)^{n-x}

distribution fn = \text{Prob}(B \leq x) = \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k}
Poisson Trial

• \( A_i \) is 1 with prob \( P_i \) and 0 with prob \( 1 - P_i \)

• Suppose \( B' \) is the sum of \( n \) independent Poisson trials

\( A_i \) with probability \( P_i \) for \( i > 1, \ldots, n \)
Hoeffding’s Theorem

• $B'$ is upper bounded by a Binomial Variable $B$

• Parameters $n, p$ where $p = \frac{\sum_{i=1}^{n} P_i}{n}$
Geometric Variable $V$

- parameter $p$

$$\forall x \geq 0 \quad \text{Prob}(V=x) = p(1-p)^x$$

procedure

GEOMETRIC parameter $p$

begin

V ← 0

loop with probability 1-p

goto exit

mean $\mu = \frac{1-p}{p}$
Probabilistic Inequalities

- For Random Variable $A$

\[
\text{mean } \mu = \bar{A} \\
\text{variance } \sigma^2 = \bar{A}^2 - (\bar{A})^2
\]
Markov and Chebychev Probabilistic Inequalities

- **Markov Inequality** (uses only mean)
  \[ \text{Prob} (A \geq x) \leq \frac{\mu}{x} \]

- **Chebychev Inequality** (uses mean and variance)
  \[ \text{Prob} \left( |A - \mu| \geq \Delta \right) \leq \frac{\sigma^2}{\Delta^2} \]
Example of use of Markov and Chebychev Probabilistic Inequalities

- If \( B \) is a Binomial with parameters \( n,p \)

\[
\text{Then } \text{Prob} \left( B \geq x \right) \leq \frac{np}{x}
\]

\[
\text{Prob} \left( \left| B - np \right| \geq \Delta \right) \leq \frac{np \left(1-p\right)}{\Delta^2}
\]
Gaussian Density Function

\[ \Psi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \]
Normal Distribution

\[ \Phi(X) = \int_{-\infty}^{x} \Psi(Y) \, dY \]

- Bounds \( x > 0 \) (Feller, p. 175)

\[ \Psi(x) \left( \frac{1}{x} - \frac{1}{x^3} \right) \leq 1 - \Phi(x) \leq \frac{\Psi(x)}{x} \]

\[ \forall x \in [0,1] \]

\[ \frac{x}{\sqrt{2\pi e}} = x \, \Psi(1) \leq \Phi(x) - \frac{1}{2} \leq \Psi(0) = \frac{x}{\sqrt{2\pi}} \]
Sums of Independently Distributed Variables

- Let $S_n$ be the sum of $n$ independently distributed variables $A_1, \ldots, A_n$

- Each with mean $\frac{\mu}{n}$ and variance $\frac{\sigma^2}{n}$

- So $S_n$ has mean $\mu$ and variance $\sigma^2$
Strong Law of Large Numbers: Limiting to Normal Distribution

- The probability density function of

\[ T_n = \frac{(S_n - \mu)}{\sigma} \]

limits as \( n \to \infty \)

to normal distribution \( \Phi(x) \)

- Hence

\[ \text{Prob} \left( \left| S_n - \mu \right| \leq \sigma x \right) \to \Phi(x) \text{ as } n \to \infty \]
Strong Law of Large Numbers (cont’d)

- So

\[
\Pr(\left| S_n - \mu \right| \geq \sigma x) \rightarrow 2(1 - \Phi(x)) \\
\leq 2\Psi(x)/x
\]

(since \(1 - \Phi(x) \leq \Psi(x)/x\))
Advanced Material
Moment Generating Functions and
Chernoff Bounds
Moments of Random Variable A (cont’d)

• $n^{th}$ Moments of Random Variable A

\[
\begin{align*}
A^n &= \int_{-\infty}^{\infty} x^n f_A(x) \, dx
\end{align*}
\]

• Moment generating function

\[
\begin{align*}
M_A(s) &= \int_{-\infty}^{\infty} e^{sx} f_A(x) \, dx \\
&= E(e^{sA})
\end{align*}
\]
Moments of Random Variable A (cont’d)

- Note
  S is a formal parameter

\[
\bar{A}^n = \left[ \frac{d^n M_A(s)}{d s^n} \right]_{s=0}
\]
Moments of Discrete Random Variable A

- $n$’th moment

\[ A^n = \sum_{x=0}^{\infty} x^n f_A(x) \]
Probability Generating Function of Discrete Random Variable $A$

$$G_A(z) = \sum_{x=0}^{\infty} z^x f_A(x) = E(z^A)$$

1st derivative $G'_A(1) = \bar{A}$

2nd derivative $G''_A(1) = \bar{A}^2 - \bar{A}$

Variance $\sigma^2_A = G''_A(1) + G'_A(1) - (G'_A(1))^2$
Moments of AND of Independent Random Variables

- If $A_1, ..., A_n$ independent with same distribution

\[ f_{A_1}(x) = f_{A_i}(x) \text{ for } i = 1, ..., n \]

Then if $B = A_1 \land A_2 \land ... \land A_n$

\[ f_B(x) = \left( f_{A_1}(x) \right)^n \]

\[ M_B(s) = \left( M_{A_1}(s) \right)^n, \quad G_B(z) = \left( G_{A_1}(z) \right)^n \]
Generating Function of Binomial Variable $B$ with Parameters $n,p$

$$G(z) = (1-p+pz)^n = \sum_{k=0}^{x} z^k \binom{n}{k} p^k (1-p)^{n-k}$$

- Interesting fact

$$\text{Prob}(B=\mu) = \Omega\left(\frac{1}{\sqrt{n}}\right)$$
Generating Function of Geometric Variable with parameter $p$

$$G(Z) = \sum_{k=0}^{\infty} Z^k (p(1-p)^k) = \frac{p}{1-(1-p)Z}$$
Chernoff Bound of Random Variable A

- Uses all moments
- Uses moment generating function

\[
\text{Prob (} A \geq x \text{) } \leq e^{-sx} M_A(s) \text{ for } s \geq 0
\]

\[
= e^{\gamma(s) - sx} \text{ where } \gamma(s) = \ln (M_A(s))
\]

\[
\leq e^{\gamma(s) - s \gamma'(s)}
\]

By setting \( x = \gamma'(s) \)

1st derivative minimizes bounds
Chernoff Bound of Discrete Random Variable $A$

\[
\text{Prob} \ (A \geq x) \leq z^{-x} \ G_A(z) \quad \text{for } z \geq 1
\]

- Choose $z = z_0$ to \textit{minimize} above bound
- Need Probability Generating function

\[
G_A(z) = \sum_{x \geq 0} z^x \ f_A(x) = E(z^A)
\]
Chernoff Bounds for Binomial B with parameters n, p

- Above mean \( x \geq \mu \)

\[
\Pr(B \geq x) \leq \left( \frac{n-\mu}{n-x} \right)^{n-x} \left( \frac{\mu}{x} \right)^x
\]

\[
\leq e^{x-\mu} \left( \frac{\mu}{x} \right)^x \text{ since } \left( 1 - \frac{1}{x} \right)^x < e^{-1}
\]

\[
\leq e^{-x-\mu} \text{ for } x \geq \mu e^2
\]
Chernoff Bounds for Binomial B with parameters n,p (cont’ d)

• Below mean $x \leq \mu$

\[
\text{Prob}(B \leq x) \leq \left( \frac{n-\mu}{n-x} \right)^{n-x} \left( \frac{\mu}{x} \right)^x
\]
Anguin-Valiant’s Bounds for Binomial B with Parameters n,p

• Just above mean

\[ \mu = np \quad \text{for} \quad 0 < \varepsilon < 1 \]

\[ \Pr(\mathcal{B} \geq (1 + \varepsilon)\mu) \leq e^{-\varepsilon^2 \frac{\mu}{2}} \]

• Just below mean

\[ \mu < np \quad \text{for} \quad 0 < \varepsilon < 1 \]

\[ \Pr(\mathcal{B} \leq \lfloor (1 - \varepsilon)\mu \rfloor) \leq e^{-\varepsilon^2 \frac{\mu}{3}} \]
Anguin-Valiant’s Bounds for Binomial $B$ with Parameters $n,p$ (cont’d)

Tails are bounded by Normal distributions
Binomial Variable $B$ with Parameters $p, N$ and Expectation $\mu = np$

- By Chernoff Bound for $p < \frac{1}{2}$
  \[
  \text{Prob} \left( B \geq \frac{N}{2} \right) < 2^N p^2
  \]

- Raghavan-Spencer bound for any $\delta > 0$
  \[
  \text{Prob} \left( B \geq (1+\delta)\mu \right) \leq \left( \frac{e^\delta}{(1+\delta)^{1+\delta}} \right)^\mu
  \]